



## PATENT ABSTRACTS OF JAPAN

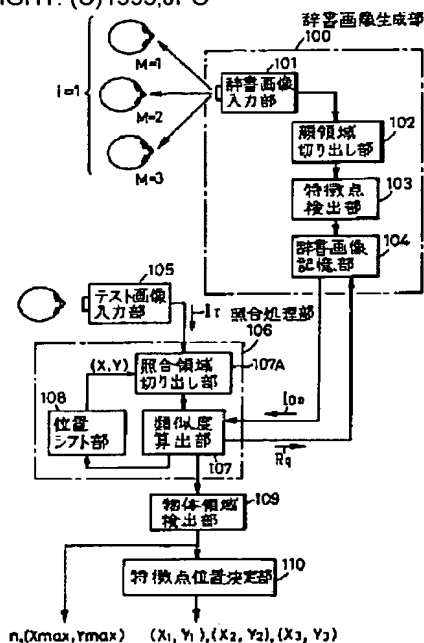
(11) Publication number: **07302327 A**(43) Date of publication of application: **14 . 11 . 95**(51) Int. Cl. **G06T 1/00**  
**G06T 7/00**(21) Application number: **06188376**(22) Date of filing: **10 . 08 . 94**(30) Priority: **11 . 08 . 93 JP 05219011**  
**13 . 10 . 93 JP 05278912**  
**09 . 03 . 94 JP 06 38005**(71) Applicant: **NIPPON TELEGR & TELEPH**  
**CORP <NTT>**(72) Inventor: **KOIKE HIDEKI**  
**SHIMADA SATOSHI**  
**TOMONO AKIRA**  
**ISHII KENICHIRO**  
**ISO TOSHIKI**(54) **METHOD AND DEVICE FOR DETECTING IMAGE**  
**OF OBJECT**

COPYRIGHT: (C)1995,JPO

## (57) Abstract:

PURPOSE: To detect an object area in a picture independently of the photographing direction of the object.

CONSTITUTION: In a dictionary image generating part 100, an object belonging to an objective category (one of L sorts) is previously photographed with a camera in a prescribed M direction, an area (collating area) of the object to be recognized is extracted from the obtained image, dictionary images ( $N=L \times M$  pieces,  $n=1, 2, \dots, N$ ) representing the direction and the category are generated and stored together with the direction (photographing direction) of the object. When a test image is given, a collating processing means 106 calculates the similarity  $r(n, X, Y)$  of a collating position  $(X, Y)$  in the test image with the  $n$ -th dictionary image in the image area and executes to collating processing while changing a collating position  $(x, y)$  and a dictionary image number  $(n)$  and an object area detecting part 109 detects and outputs a collating position  $(X_{\max}, Y_{\max})$  obtained at the time of the highest similarity  $(r)$ .



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-302327

(43) 公開日 平成7年(1995)11月14日

(51) Int.Cl.<sup>6</sup>

識別記号

庁内整理番号

F I

技術表示箇所

G 0 6 T 1/00

7/00

9061-5L

G.0 6 F 15/ 62

15/ 70

3 8 0

4.5 5 A

審査請求 未請求 請求項の数29 O L (全 17 頁)

(21) 出願番号 特願平6-188376

(22) 出願日 平成6年(1994)8月10日

(31) 優先権主張番号 特願平5-219011

(32) 優先日 平5(1993)8月11日

(33) 優先権主張国 日本 (J P)

(31) 優先権主張番号 特願平5-278912

(32) 優先日 平5(1993)10月13日

(33) 優先権主張国 日本 (J P)

(31) 優先権主張番号 特願平6-38005

(32) 優先日 平6(1994)3月9日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区内幸町一丁目1番6号

(72) 発明者 小池 秀樹

東京都千代田区内幸町1丁目1番6号 日

本電信電話株式会社内

(72) 発明者 高田 聡

東京都千代田区内幸町1丁目1番6号 日

本電信電話株式会社内

(72) 発明者 伴野 明

東京都千代田区内幸町1丁目1番6号 日

本電信電話株式会社内

(74) 代理人 弁理士 草野 卓

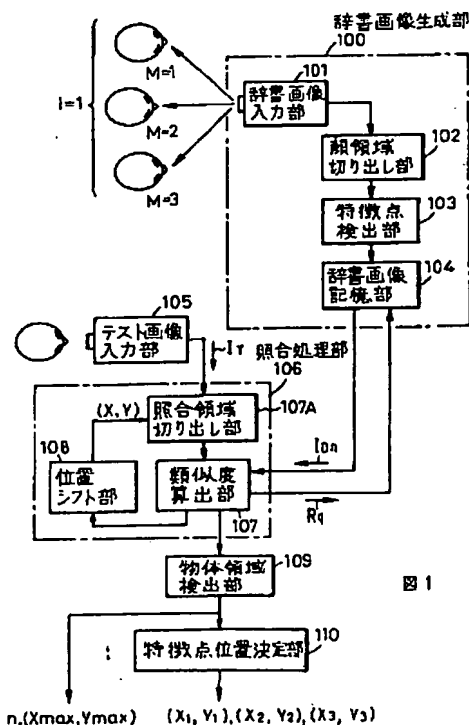
最終頁に続く

(54) 【発明の名称】 物体画像検出方法及び検出装置

(57) 【要約】

【目的】 物体の撮影方向によらず、画像中の物体領域を検出可能にする。

【構成】 辞書画像生成部において対象となるカテゴリー (L種類) に属する物体をあらかじめ所定のM方向からカメラで撮影し、得られた画像から認識対象物体の領域 (照合領域) を抽出しその方向とそのカテゴリーを代表する辞書画像 ( $N=L \times M$  個,  $n=1, 2, \dots, N$ ) を生成し、それら辞書画像を、物体の向き (撮影方向) とともに記憶しておく。テスト画像が与えられると、照合処理部によりそのテスト画像中の照合位置 (X,Y) の画像領域に対してn番目の辞書画像との類似度  $r(n, X, Y)$  を算出し、その照合処理を照合位置 (x,y) 及び辞書画像番号 n を変えて実行し、物体領域検出部により類似度 r が最も高くなるときの照合位置 ( $X_{\max}, Y_{\max}$ ) を検出し出力する。



Best Available Copy

## 【特許請求の範囲】

【請求項1】 L種類のカテゴリーの参照物体のそれぞれを予め決めたM通りの方向から撮影して得た画像から生成された辞書画像  $I_{Dn}$ ,  $n=1, 2, \dots, N$ ,  $N=L \times M$ ,  $L$ は1以上の整数、 $M$ は2以上の整数、を使って入力テスト画像中の所望の物体画像を物体領域として検出する物体画像検出方法であり、以下のステップを含む：

- (a) 上記入力テスト画像に対する照合位置  $(X, Y)$  を更新して出力し、
- (b) 上記入力テスト画像中の上記照合位置  $(X, Y)$  を基準として所定の大きさの領域の画像を照合領域画像として切り出し、
- (c) 上記照合領域画像と第  $n$  番の上記辞書画像との類似度  $r(n)$  を計算し、
- (d) 上記照合位置  $(X, Y)$  と上記  $n$  を変えて上記類似度  $r(n)$  を計算することを繰り返し、それによって  $L$  カテゴリー、 $M$  方向のそれぞれの上記辞書画像との類似度  $r(n)$  を得て、
- (e) 上記ステップ(d) によって得られた上記類似度が最大となるときの上記照合位置を物体位置  $(X_{\max}, Y_{\max})$  として得ることにより、上記テスト画像中の上記物体領域を検出する。

【請求項2】 請求項1記載の物体画像検出方法において、上記テスト画像中の上記物体と、着目した少なくとも1つの上記カテゴリーを代表する辞書画像との類似度  $r(n)$  の大きさにより、上記テスト画像中の上記物体が上記着目したカテゴリーに属するか否かを判別するステップを含む。

【請求項3】 請求項2記載の物体画像検出方法において、上記判別ステップは上記類似度  $r(n)$  を予め決めた閾値と比較するステップと、その比較結果に基づいて上記テスト画像中の上記物体が予め決めた少なくとも1つのカテゴリーに属するか否かを決定するステップを含む。

【請求項4】 請求項1又は2記載の物体画像検出方法において、上記  $M$  は隣接する方向角が  $10$  度以上  $40$  度以下になるように選択される。

【請求項5】 請求項1又は2記載の物体画像検出方法において、上記辞書画像は、上記参照物体の画像の中の照合領域を縦方向に  $w$  個及び横方向に  $v$  個に分割して得られる  $v \times w = q$  個の各ブロックの代表画素値で構成される  $q$  個のブロック画像情報からなる辞書ブロック画像情報列として生成し、  
上記類似度を計算するステップは、上記テスト画像中の上記照合位置  $(X, Y)$  における上記照合領域を縦方向に  $w$  個及び横方向に  $v$  個に分割して得られる  $v \times w = q$  個の各ブロックの代表画素値で構成される  $q$  個のブロック画像情報からなるテストブロック画像情報列を生成し、 $n$  番目の上記辞書ブロック画像情報列と上記テストブロック画像情報列との類似度を上記類似度  $r(n)$  として算出するステップである。

【請求項6】 請求項5記載の物体画像検出方法において、上記辞書画像との最大の上記類似度  $r(n)$  を与えるカテゴリーを選択し、それにより上記テスト画像中の物体が属するカテゴリーを判別するステップを含む。

【請求項7】 請求項5記載の物体画像検出方法において、 $v \times w = q$  個の各上記ブロックを構成する画素数は予め決めた範囲内に限定されている。

【請求項8】 請求項1又は2記載の物体画像検出方法において、上記類似度を得るステップは、上記照合領域画像と上記辞書画像のそれぞれの照合領域内の画素の濃度の平均と分散がそれぞれ予め決めた値になるように、上記照合領域画像の各画素の濃度を変換し、濃度を変換された上記照合領域画像の辞書画像に対する類似度を得るステップである。

【請求項9】 請求項1または8記載の物体画像検出方法において、最大の上記類似度を与えるカテゴリーを選択し、それにより上記テスト画像中の物体が属するカテゴリーを判別するステップを含む。

【請求項10】 請求項1または2記載の物体画像検出方法において、 $L$  カテゴリーに属する上記参照物体を上記  $M$  方向から撮影して得た画像から上記参照物体上の予め決めた照合部の画像を上記辞書画像として切り出し、上記参照物体の予め決めた少なくとも1つの特徴点に対応する各上記辞書画像中の特徴点  $P_f$  の位置  $(x_f, y_f)$  を上記辞書画像と対応して辞書画像記憶手段に記憶し、上記テスト画像中の上記物体位置  $(X_{\max}, Y_{\max})$  に類似度が最も高い辞書画像を重ねて、該辞書画像に対応する上記特徴点の位置を上記テスト画像中の位置  $(X_f, Y_f)$  に変換して出力する、ステップを含む。

【請求項11】 請求項10記載の物体画像検出方法において、上記辞書画像  $I_{Dn}$ ,  $n=1, 2, \dots, N$ , の各々において上記特徴点  $P_f$  を含む所定領域から部分辞書画像を生成し、上記部分辞書画像を上記辞書画像の番号  $n$  および上記位置  $(x_f, y_f)$  とともに部分辞書手段に記憶しておく、  
上記テスト画像が与えられると、上記特徴点  $P_f$  の位置  $(X_f, Y_f)$  の周辺において、上記  $n$  番目の辞書画像から作られた上記部分辞書画像との類似度  $r_f$  を部分照合処理として算出し、  
上記部分照合処理を照合位置  $(X, Y)$  を変えて実行し類似度  $r_f$  が最も高くなるときの位置を特徴点  $P_f$  の最適位置  $(X_{f\max}, Y_{f\max})$  とし、上記最適位置を各特徴点毎に出力するステップを含む。

【請求項12】 請求項10記載の物体画像検出方法において、  
上記テスト画像における認識対象物体の領域を照合領域として切り出し、  
基準となる方向から上記参照物体を撮影したときの画像から抽出した特徴点の間の位置関係を位置関係記憶手段に記憶し、

上記辞書画像および上記テスト画像の照合領域からそれぞれ特徴点を抽出し、

上記辞書画像および上記テスト画像の特徴点の位置関係が上記位置関係記憶手段で記憶している関係になるように、上記 $n$ 番目の辞書画像および上記テスト画像の照合領域を幾何学的に変換し、

上記テスト画像と上記 $n$ 番目の辞書画像の上記幾何学的に変換された照合領域間の類似度 $r(n)$ を算出する、ステップを更に含む。

【請求項13】 請求項1記載の物体画像検出方法において、上記辞書画像と上記類似度 $r(n)$ の大きさと、上記部分辞書画像と上記類似度 $r_f$ の大きさから、上記テスト画像中の物体が対象とするどのカテゴリーに属するかを判別するステップを含む。

【請求項14】 請求項4記載の物体画像検出方法において、上記テストブロック画像情報列を生成するステップは、上記入力テスト画像を得る所定時間前に同じ方向から撮影して得た入力画像を背景画像として保持するステップと、現在の上記入力テスト画像と上記背景画像との差の画像を生成するステップと、上記差の画像から物体領域の大きさを求めるステップと、上記物体領域の大きさに基づいて上記ブロックサイズの範囲を決定するステップと、上記類似度 $r(n)$ を算出する処理を上記ブロックサイズの範囲で初期値から一定幅ずつブロックサイズを変更して繰り返すステップとを含む。

【請求項15】  $L$ 種類のカテゴリーの参照物体のそれぞれを予め決めた $M$ 通りの方向から撮影して得た画像から生成された辞書画像 $I_{Dn}$ 、 $n=1,2,\dots,N$ 、 $N=L \times M$ 、 $L$ は1以上の整数、 $M$ は2以上の整数、を使って入力テスト画像中の所望の物体画像を物体領域として検出する物体画像検出方法であり、以下のステップを含む：

(a) 入力テスト画像に対する照合位置 $(X,Y)$ を更新して出力し、

(b) 上記入力テスト画像中の上記照合位置 $(X,Y)$ を基準として所定の大きさの領域の画像を照合領域画像として切り出し、

(c) 上記辞書画像および上記テスト画像の照合領域からそれぞれ特徴点を抽出し、

(d) 基準となる方向から上記参照物体を撮影したときの画像から抽出した特徴点の間の位置関係を位置関係記憶手段に記憶し、

(e) 上記辞書画像および上記テスト画像の特徴点の位置関係が上記位置関係記憶手段で記憶している関係になるように、上記 $n$ 番目の辞書画像および上記テスト画像の照合領域を幾何学的に変換し、

(f) 上記テスト画像と上記 $n$ 番目の辞書画像の上記幾何学的に変換された照合領域間の類似度 $r(n)$ を算出し、

(g) 上記類似度 $r(n)$ の大きさから上記テスト画像中の物体が属するカテゴリーを判別する、ステップを含む。

【請求項16】  $L$ 種類のカテゴリーの参照物体のそれ

それを予め決めた $M$ 通りの方向から撮影して得た画像から生成された辞書画像 $I_{Dn}$ 、 $n=1,2,\dots,N$ 、 $N=L \times M$ 、 $L$ は1以上の整数、 $M$ は2以上の整数、を記憶する辞書画像記憶手段と、

上記入力テスト画像に対する照合位置 $(X,Y)$ を更新して出力する位置シフト手段と、

上記入力テスト画像中の上記照合位置 $(X,Y)$ を基準として所定の大きさの領域の画像を照合領域画像として切り出す照合領域切り出し手段と、

上記照合領域画像と第 $n$ 番の上記辞書画像との類似度 $r(n)$ をそれぞれ計算する類似度算出手段と、

上記照合位置 $(X,Y)$ と上記 $n$ を変えて上記類似度 $r(n)$ を計算することを繰り返し、それによって得られた $L$ カテゴリー、 $M$ 方向のそれぞれの上記辞書画像との類似度 $r(n)$ が与えられ、上記類似度が最大となる時の上記照合位置を物体位置 $(X_{\max}, Y_{\max})$ として得ることにより、上記テスト画像中の上記物体領域を物体画像として検出する物体領域検出手段、とを含む物体画像検出装置。

【請求項17】 請求項16記載の物体画像検出装置において、上記テスト画像中の物体と対象となるカテゴリーを代表する辞書画像と上記類似度 $r(n)$ の大きさにより、上記テスト画像中の物体が対象となる予め決めた少なくとも1つの上記カテゴリーに属するか否かを判別する手段を更に含む。

【請求項18】 請求項17記載の物体画像検出装置において、上記判別する手段は上記類似度 $r(n)$ を予め決めた閾値と比較し、その比較結果に基づいて上記テスト画像中の上記物体が上記予め決めた少なくとも1つのカテゴリーに属するか否かを決定する手段である。

【請求項19】 請求項16記載の物体画像検出装置において、 $L$ カテゴリーに属する上記参照物体を上記 $M$ 方向から撮影して得た画像から上記参照物体上の予め決めた照合部の画像を上記辞書画像として切り出す辞書画像切り出し手段と、

上記参照物体の予め決めた $F$ 個、 $F$ は1以上の整数、の特徴点に対応する各上記辞書画像中の特徴点 $P_f$ の位置 $(x_f, y_f)$ を上記辞書画像と対応して記憶する辞書画像記憶手段と、

上記テスト画像中の上記物体位置 $(X_{\max}, Y_{\max})$ に類似度が最も高い辞書画像を重ねて、上記辞書画像に対応する上記特徴点の位置を上記テスト画像中の位置 $(x_f, y_f)$ に変換して出力する特徴点位置検出手段、を含む。

【請求項20】 請求項16又は17記載の物体画像検出装置において、上記 $M$ は隣接する方向角が10度以上40度未満になるように選択される。

【請求項21】 請求項16又は17記載の物体画像検出装置において、上記辞書画像は、上記参照物体の画像の中の照合領域を縦方向に $w$ 個及び横方向に $v$ 個に分割して得られる $v \times w = q$ 個の各ブロックの代表画素値で構成される $q$ 個のブロック画像情報からなるブロック画

像情報列として上記辞書画像記憶手段に記憶されており、

上記類似度算出手段は、上記テスト画像中の上記照合位置 $(X, Y)$ における上記照合領域を縦方向に $w$ 個及び横方向に $v$ 個に分割して得られる $v \times w = q$ 個の各ブロックの代表画素値で構成される $q$ 個のブロック画像情報からなるブロック画像情報列を生成し、 $n$ 番目の上記辞書画像との類似度を上記類似度 $r(n)$ として算出する手段である。

【請求項22】 請求項21記載の物体画像検出装置において、 $v \times w = q$ 個の各上記ブロックを構成する画素数は予め決めた範囲内に限定されている。

【請求項23】 請求項16又は17記載の物体画像検出装置において、上記照合領域内の各画素の濃度の平均と分散が一定になるように、各画素の濃度を変換する明るさ正規化手段を更に含む。

【請求項24】 請求項16、21または23記載の物体画像検出装置において、出力される最大の上記類似度 $r(n)$ を与えるカテゴリーを選択し、それにより上記テスト画像中の物体が属するカテゴリーを判別する判別処理手段を更に含む。

【請求項25】 請求項17記載の物体画像検出装置において、上記参照物体の予め決めた $F$ 個、 $F$ は1以上の整数、の特徴点に対応する各上記辞書画像中の特徴点 $P_f$ の位置 $(x_f, y_f)$ を上記辞書画像と対応して記憶する辞書画像記憶手段と、

上記テスト画像中の上記物体位置 $(X_{\max}, Y_{\max})$ に類似度が最も高い辞書画像を重ねて、該辞書画像に対応する上記特徴点の位置を上記テスト画像中の位置 $(X_f, Y_f)$ に変換して出力する特徴点位置検出手段と、を更に含む。

【請求項26】 請求項18または25記載の物体画像検出装置において、上記辞書画像 $I_{Dn}$ 、 $n=1, 2, \dots, N$ 、の各々において上記特徴点 $P_f$ を含む所定領域から部分辞書画像を生成し、上記部分辞書画像を上記辞書画像の番号 $n$ および上記位置 $(x_f, y_f)$ とともに記憶する部分辞書手段と、

上記テスト画像が与えられると、上記特徴点 $P_f$ の位置 $(X_f, Y_f)$ の周辺において、上記 $n$ 番目の辞書画像から作られた上記部分辞書画像との類似度 $r_f$ を部分照合処理として算出する部分照合処理手段と、

上記部分照合処理を照合位置 $(X, Y)$ を変えて実行し類似度 $r_f$ が最も高くなる時の位置を特徴点 $P_f$ の最適位置 $(X_{f\max}, Y_{f\max})$ とし、上記最適位置を各特徴点毎に出力する特徴点位置高精度検出手段と、を更に含む。

【請求項27】 請求項26記載の物体画像検出装置において、上記辞書画像との上記類似度 $r(n)$ の大きさと、上記部分辞書画像との上記類似度 $r_f$ の大きさから、上記テスト画像中の物体が対象とするどのカテゴリーに属するかを判別する判別処理手段を含む。

【請求項28】 請求項25記載の物体画像検出装置に

おいて、

上記テスト画像における認識対象物体の領域を照合領域として切り出す照合領域切り出し手段と、

基準となる方向から上記参照物体を撮影したときの画像から抽出した特徴点の間の位置関係を記憶する位置関係記憶手段と、

上記辞書画像および上記テスト画像の照合領域からそれぞれ特徴点を抽出する特徴点検出手段と、

上記辞書画像および上記テスト画像の特徴点の位置関係が上記位置関係記憶手段で記憶している関係になるように、上記 $n$ 番目の辞書画像および上記テスト画像の照合領域を幾何学的に変換する幾何学的正規化手段と、を更に含む。

【請求項29】  $L$ 種類、 $L$ は1以上の整数、のカテゴリの参照物体のそれぞれを予め決めた $M$ 通り、 $M$ は2以上の整数、の方向から撮影して得た画像から生成された辞書画像 $I_{Dn}$ 、 $n=1, 2, \dots, N$ 、 $N=L \times M$ 、を使って入力テスト画像中の所望の物体領域を物体画像として検出する物体画像検出装置であり以下を含む：上入力テスト画像に対する照合位置 $(X, Y)$ を更新して出力する位置シフト手段と、

上記入力テスト画像中の上記照合位置 $(X, Y)$ を基準として所定の大きさの領域の画像を照合領域画像として切り出す照合領域切り出し手段と、

上記辞書画像および上記テスト画像の照合領域からそれぞれ特徴点を抽出する特徴点検出手段と、

基準となる方向から上記参照物体を撮影したときの画像から抽出した特徴点の間の位置関係を記憶する位置関係記憶手段と、

上記辞書画像および上記テスト画像の特徴点の位置関係が上記位置関係記憶手段で記憶している関係になるように、上記 $n$ 番目の辞書画像および上記テスト画像の照合領域を幾何学的に変換する幾何学的正規化手段と、

上記テスト画像と上記 $n$ 番目の辞書画像の上記幾何学的に変換された照合領域間の類似度 $r(n)$ を算出する照合処理手段と、及び上記類似度 $r(n)$ の大きさから上記テスト画像中の物体が属するカテゴリーを判別する判別処理手段。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は、人物顔など、特徴がある物体を任意の方向からカメラで撮影し得られた画像をあらかじめ生成した辞書画像と照合処理することによりその画像中の該物体領域を検出する方法、および、該検出領域の物体がカテゴリーに分類されて登録済みの人物顔など、あらかじめ定めたカテゴリーのどれに属するか、および、予め定めたカテゴリーに属するか否かを検出する方法及びそれを使った装置に関するものである。

【0002】

【従来の技術】例えば従業員や顧客のように予め決めた

特定の人間のみ特定の部屋、建物、あるいは区域への入出を許可するため、入出しようとする人を自動的に識別することが必要とされる場合がある。あるいは銀行に入入りする不特定多数の人間の中から不審な人物を判別することが必要とされる場合がある。このような場合、特定の場所に入出する人間又はしようとする人間をカメラで撮影し、その画像からその人物の顔を識別又は判別することが提案されている。

【0003】従来の画像中物体領域抽出装置には、濃淡情報や色彩情報の閾値処理に基づいた方法がある。例えば、画像から人物の顔領域を抽出するときは、肌に相当する濃度値や彩度・色相を求めて、肌領域を検出するための閾値を設定し、画像全体を閾値処理する。また、別の方法として、物体の形状をモデル化し、画像のエッジとフィッティングさせることで物体領域を抽出する方法がある。

【0004】従来の物体識別装置には、例えば、対象となる人物の正面顔又は横顔など特定方向の顔の画像を辞書に用意しておき、テスト画像との照合をとって、人物の同定を行う装置がある（例えば Ashoc Samal, Prasa na A. Iyengar: "Automatic recognition and analysis of human faces and facial expressions" Pattern recognition, Vol.25, No.1, pp65-77, 1992）。

【0005】

【発明が解決しようとする課題】従来の閾値処理による物体抽出装置では、最適な閾値を設定する必要がある。この最適な閾値は、物体の向きによって変動することが多い。従って、物体の向きが未知の場合には、物体領域を抽出するための閾値を設定することが困難であった。また、形状モデルとエッジをフィッティングさせる方法では、物体の向きにより画像での物体の形状が変化したり、物体の一部の隠蔽や陰影などにより、物体の輪郭をあらわすエッジを画像から正しく抽出できなかった場合には、精度よく物体領域を抽出することができなかった。

【0006】従来の物体識別装置では、辞書に用意しておく画像と同じ方向から撮影した画像（正面顔）を入力する必要があり、物体の撮影方向を何らかの方法で限定するために、適用できる範囲が限られていた。本発明の目的は、物体の撮影方向によらず、画像中の該物体領域を高速な処理で安定に検出できる方法と装置、及び物体があらかじめ定めたカテゴリー（登録してある人物、子供顔、成人女性顔、成人男性顔など）のどれに属するか検出する方法と装置、及び検出された物体があらかじめ定めたカテゴリー（人物顔、子供顔、成人女性顔、成人男性顔）に属しているか否かを検出する方法と装置を提供することにある。

【0007】

【課題を解決するための手段】この発明によれば、対象となるカテゴリー（L種類）に属する物体を予め複数の

所定のM方向からカメラで撮影し、得られた画像から認識対象物体の領域（照合領域）を抽出し該方向と該カテゴリを代表する辞書画像（ $N = M \times L$ 個、 $n = 1, 2, \dots, N$ ）を生成し、辞書画像生成手段にそれら辞書画像を物体の向き（撮影方向）とともに記憶しておく。テスト画像が与えられると、そのテスト画像中の照合位置（ $X, Y$ ）に対して前記n番目の辞書画像との類似度 $r(n, X, Y)$ を照合処理手段により算出し、その照合処理を（ $X, Y$ ）、および、nを変えて実行し、類似度rが最も高くなるときの位置（ $X_{\max}, Y_{\max}$ ）を物体領域に対応する領域として出力する。

【0008】この発明による物体領域検出方法及び装置によれば、辞書画像生成手段において、対象となるカテゴリー（L種類）に属する物体を代表する辞書画像の方向数Mを適切に設定しておけば、どんな方向を向いた認識対象物体を含むテスト画像が与えられても、該テスト画像の照合領域と辞書画像との類似度が最大になる該テスト画像の照合領域の位置が該認識対象物体の位置に一致し、そのときの辞書画像が該認識対象物体が属すカテゴリを代表するものとして行うことができる。従って、テスト画像における認識対象物体の向きによらずに該認識対象物体を含むように照合領域を抽出することができる。

【0009】

【実施例】以下に説明する各実施例では、対象となるカテゴリーiとして、子供顔（ $i=1$ ）、成人女顔（ $i=2$ ）、成人男顔（ $i=3$ ）を、領域抽出の対象として顔領域を、顔領域内のK個の特徴点として左右の目と口の3点を選んだ場合を例に説明するが、顔以外の物体を識別する場合にも適用できることは容易に理解できよう。

【0010】図1はこの発明による第1の実施例の処理と構成を示す機能ブロック図である。図1において、100は辞書画像生成部であって、辞書画像入力部101、顔領域切り出し部102、特徴点検出部103、辞書画像記憶部104より構成される。辞書画像入力部101はカメラにより構成され、 $i = 1, 2, 3$ の何れかのカテゴリーに属する合計p人のそれぞれについてM方向の顔画像を取り込み、得られた画像を顔領域切り出し部102に出力する。

【0011】顔領域切り出し部102は1フレームの画像を保持する画像バッファを有し、受けた各画像から顔領域を切り出し、予め決めた一定の大きさの顔領域のみの画像として出力する。ここでは、辞書画像を生成するための処理であるので、人手を介して顔領域を切り出しても実用上問題にならない。特徴点検出部103は顔領域切り出し部102より受けた顔画像から予め決めた特徴項目である例えば目と口の特徴点を検出し、それらの特徴点の位置を顔画像と共に辞書画像記憶部104に供給する。ここでの特徴点検出でも、辞書画像を生成するための処理であるので、人手を介して検出しても実用上問題にならない。辞書画像記憶部104は特徴点検出部103より受

けた特徴点の位置と顔画像から辞書画像を生成して記憶する。

【0012】辞書画像の生成は、例えば、各カテゴリー  $i=1, 2, 3$  に属する  $p_1, p_2, p_3$  人、合計  $p$  人のそれぞれについて予め決めた  $M$  方向の顔画像を、それぞれの方向の顔画像毎に目と口の位置を用いて顔領域の大きさと位置を合わせてから平均することにより得られる。あるいは、各カテゴリー毎に、 $p \times M$  個の顔画像を  $M$  個のクラスタに分け、各画像が  $a \times b$  画素から成るとすると、これら全画素を要素とする  $a \times b$  次元のベクトル（画像ベクトルと呼ぶ）でその画像を表現した場合、各クラスタ内の全画像ベクトルの重心ベクトルであるセントロイド、または、セントロイドに最も近い画像ベクトル（顔画像）を辞書画像としてもよい。この様にして、例えば図2に示すように、カテゴリー  $i=1, 2, \dots, L$  のそれぞれについて  $M$  個の方向から見た代表的な顔画像と、それらの各画像中の3つの特徴点の位置座標  $(x_1, y_1), (x_2, y_2), (x_3, y_3)$  が対応づけられた辞書が得られる。図2ではこれらの特徴点の位置座標を  $(x_f, y_f)$  で示している。辞書画像記憶部104は照合処理部106から辞書画像の読み出し要求を受けると、記憶している辞書画像を読みだし出力する。

【0013】カメラにより構成されたテスト画像入力部105は検査対象の画像を取り込み、取り込んだ画像をテスト画像  $I_T$  として照合処理部106に出力する。照合処理部106は照合領域切り出し部107Aと、類似度算出部107と、位置シフト部108とにより構成される。照合領域切り出し部107Aは1フレームの画像を保持する画像バッファを有し、テスト画像入力部105からテスト画像  $I_T$  を受けると、位置シフト部108から指定された照合位置初期値  $(X_0, Y_0)$  を基準とする所定の大きさの照合領域を切り出し、類似度算出部107に与える。類似度算出部107は照合領域画像を受けると、辞書画像記憶部104に辞書画像の読み出し要求を出す。辞書画像記憶部104は類似度算出部107の要求に回答して図2に示す  $N$  個の辞書画像  $I_{Dn} (n=1, 2, \dots, N)$  を特徴点位置  $(x_f, y_f)$  と共に順次読みだして類似度算出部107に出力する。次に類似度算出部107はテスト画像の照合領域と、 $N$  個の辞書画像との類似度  $r(n, X_0, Y_0)$ 、ただし  $n=1, 2, \dots, N$ 、をそれぞれ算出し、それらの類似度を物体領域検出部109に与えると共に、照合位置の更新命令を位置シフト部108に与える。図3は、照合位置  $(X_0, Y_0)$  でのテスト画像  $I_T$  の領域と  $n=1$  の辞書画像  $I_{D1}$  との類似度を算出する様子を示している。

【0014】照合領域切り出し部107Aは位置シフト部108より更新された照合位置  $(X, Y)$  を受け取ると、その位置を基準とする所定の大きさの照合領域をテスト画像から切り出して類似度算出部107に与える。類似度算出部107は上述と同様に、切り出されたその照合領域と辞書画像との類似度  $r(n, X, Y)$  をそれぞれ算出して物体領域検

出部109に出力するとともに、照合位置更新命令を位置シフト部108に出力する。但し、照合位置  $(X, Y)$  が予め設定した最終位置  $(X_m, Y_m)$  に到達したならば、照合処理部106の処理を終了し、物体領域検出部109の処理に進む。類似度算出部107における類似度  $r$  の計算は、例えば、2つの画像の相関係数により算出すればよい。

【0015】位置シフト部108は類似度算出部107より照合位置更新命令を受け取ると、前回の照合位置座標  $X, Y$  のいずれか一方又は両方の値をそれぞれ予め決めた歩進幅  $\Delta X, \Delta Y$  だけ変化させ、変化後の位置を新たな照合位置座標  $(X, Y)$  として類似度算出部107に出力する。位置の変化のさせ方は、例えば、テスト画像をラスタースキャンするように行えばよい。即ち、例えば  $X$  をステップ  $\Delta X$  づつ歩進し  $X_0$  から  $X_m$  までスキャンする毎に  $Y$  を  $\Delta Y$  だけ歩進する事を繰り返し、 $Y$  を  $Y_0$  から  $Y_m$  まで変化させる。

【0016】物体領域検出部109は類似度算出部107より受け取った類似度  $r(n, X, Y)$  が最大となるときに位置  $(X_{\max}, Y_{\max})$  と辞書画像  $I_{Dn}$  を検出し、特徴点位置決定部110に出力する。特徴点位置決定部110は物体領域検出部109より受けた辞書画像  $n$  をテスト画像中の照合位置  $(X_{\max}, Y_{\max})$  の領域に重ね、このときに、辞書画像に記憶されている特徴点位置  $(x_f, y_f)$  である目と口の位置  $(x_1, y_1), (x_2, y_2), (x_3, y_3)$  からテスト画像における目の位置  $(X_1, Y_1) = (X_{\max} + x_1, Y_{\max} + y_1)$ 、 $(X_2, Y_2) = (X_{\max} + x_2, Y_{\max} + y_2)$  と口の位置  $(X_3, Y_3) = (X_{\max} + x_3, Y_{\max} + y_3)$  を検出する。

【0017】位置シフト部108における更新照合位置  $(X, Y)$  が予め設定した最終照合位置座標  $(X_m, Y_m)$  に達したときには、照合処理を終了し、物体領域検出部109の処理に進む。物体領域検出部109では、受け取った類似度  $r(n, X, Y)$  のなかの最大値を検出し、このときの照合位置  $(X_{\max}, Y_{\max})$  と辞書画像  $I_{Dn}$  の番号  $n$ 、及びその画像の有する特徴点位置  $(x_f, y_f)$  を物体位置決定部110に出力する。図3に、類似度が最大となるときに位置  $(X_{\max}, Y_{\max})$  を示す。物体位置決定部110は、物体領域検出部109より受けた情報  $n, (X_{\max}, Y_{\max}), (x_f, y_f)$  からテスト画像における目と口の位置を決定する。図3に目と口の位置

$$(X_1, Y_1) = (X_{\max} + x_1, Y_{\max} + y_1)$$

$$(X_2, Y_2) = (X_{\max} + x_2, Y_{\max} + y_2)$$

$$(X_3, Y_3) = (X_{\max} + x_3, Y_{\max} + y_3)$$

を示す。テスト画像上におけるこれらの特徴点の座標を  $(X_f, Y_f); f=1, 2, 3$  と表すものとする。

【0018】また、辞書画像の物体の方向  $M$  については、 $M$  を大きくすればするほど、テスト画像  $I_T$  における物体像の向きに対する適用能力が向上する。しかし、 $M$  を大きくすると、辞書画像を記憶するために必要な辞書画像記憶部104の記憶容量が増大するとともに、照合処理106での処理に時間がかかる。図4は、図5A及び

5 Bに示すように、顔方向を真正面から左右に60度、上下に30度の範囲で、左右、上下に $\theta$ (=5,10,15,20,25,30,35,40,45,50,55,60)度の間隔で変化させた画像を $p=10$ 人について撮影して、方向別に平均した画像を辞書画像とし、辞書画像の作成に用いた人物とは別の10人について、それぞれ顔方向が真正面から左右に60度、上下に30度の範囲から選択した45方向の顔画像を取り込んだ450枚のテスト画像から顔領域を抽出する実験を行った結果を示し、縦軸は抽出成功率、横軸は隣接方向角 $\theta$ である。図4より、 $\theta$ が10度より小さくなると辞書画像の総数が急激に増え、 $\theta$ が40度以上になると顔領域の抽出精度が低下することが分かる。

【0019】そこで、この発明の1実施例では、隣接する方向角が10度以上40度未満になるように各カテゴリ毎にM個の画像を取り込み、特徴点座標 $(x_f, y_f)$ と共に辞書画像として辞書画像記憶部104に記憶することにより、比較的少ない辞書画像の数Nでも高い成功率で物体領域を抽出でき、従って、物体領域を高速に精度よく抽出できることが理解できる。

【0020】以上説明した通り、本発明によれば、顔の向きによらずに、画像における顔領域を抽出すること、および、顔領域の特徴点として目や口の位置を検出することができる。図6はこの発明による第2の実施例の処理と構成を示す機能ブロック図である。ただし、ブロック17で示す入力テスト画像と、図1の実施例において最大類似度を与える照合位置 $(X_{\max}, Y_{\max})$ から決定された特徴点の座標 $(X_f, Y_f)$ ,  $f=1, 2, 3$ と、その時の辞書画像番号 $n$ とが部分照合処理部502に入力されることを模式的に示している。部分照合処理部502は照合領域切り出し部503Aと類似度算出部503と位置シフト部504とを有し、これらの動作は図1における照合処理部106の照合領域切り出し部107A、類似度算出部107及び位置シフト部108の動作と同様である。

【0021】図6の実施例において、部分辞書画像記憶部501には図7に示すように、図1の画像記憶部104に記憶してあるカテゴリ $i$ 、顔方向Mの各顔画像 $I_{Dn}$ に対応して、その辞書画像の目・口の特徴点の座標 $(x_f, y_f)$ を含む所定の大きさの部分領域画像 $A_1, A_2, A_3$ の組が部分辞書画像 $I_{Pn}$ として記憶されている。部分照合処理部502は、図1で説明した照合処理により得られたテスト画像と最も類似度の高い辞書画像の番号 $n$ とテスト画像における特徴点位置 $(X_f, Y_f)$ である目の位置 $(X_1, Y_1), (X_2, Y_2)$ と口の位置 $(X_3, Y_3)$ が与えられると、部分辞書画像記憶部501に $n$ 番目の辞書画像 $I_{Dn}$ に対応した部分辞書画像 $I_{Pn}$ の読み取り要求を出し、部分辞書画像 $I_{Pn}$ を受け取る。部分照合処理部502は目・口の特徴点 $f$ (=1, 2, 3)について、例えば図8に $f=1$ の場合を示すように位置 $(X_f, Y_f)$ の周辺で予め設定しておいた範囲内で図1で説明した照合処理部104の処理と同様に照合位置 $(X, Y)$ を順次変えてテスト画像と部分辞書画像と

の類似度 $r_f(n, f, X, Y)$ を算出し、算出した類似度を特徴点位置高精度決定部505に出力する。特徴点位置高精度決定部505は特徴点 $f=1, 2, 3$ のそれぞれについて、類似度が最大となるときのテスト画像におけるそれぞれの照合位置 $(X_{f\max}, Y_{f\max})$ を検出し、更にこれらの照合位置と辞書画像中の特徴点 $(x_f, y_f)$ からテスト画像中の特徴点位置 $(X_{f\max}+x_f, Y_{f\max}+y_f)$ を $f=1, 2, 3$ についてそれぞれ計算する。

【0022】このように、この実施例によれば、目や口の周辺の部分領域辞書画像との照合を行うので、テスト画像における目や口などの特徴点の位置を精度よく抽出することができる。図9は図1の実施例において更に判定処理部901を設け、テスト画像中の検出された物体領域の画像が属するカテゴリを判定するようにした実施例である。ただし、この実施例においては図1における特徴点位置決定部110を必要としない。

【0023】物体領域検出部109は、図1で説明したようにしてテスト画像中の辞書画像と最も類似するときの照合位置 $(X_{\max}, Y_{\max})$ を検出すると、その位置 $(X_{\max}, Y_{\max})$ と、そのときの辞書画像 $I_{Dn}$ の番号 $n$ とを判定処理部901に出力する。判定処理部901は物体領域検出部109より受けた照合位置 $(X_{\max}, Y_{\max})$ でのテスト画像の部分領域と $n$ 番目の辞書画像 $I_{Dn}$ との類似度 $r(n, X_{\max}, Y_{\max})$ から、テスト画像における物体が辞書画像の物体のカテゴリ $i=1, 2, \dots, L$ のどれに属するかを判別し、判別結果を出力する。

【0024】判定処理部901の処理を、画像中の物体が人物の顔である場合に男性的な顔( $i=1$ )、女性的な顔( $i=2$ )、中性的な顔( $i=3$ )のどれに近いかを判別する場合を例に説明する。この場合、カテゴリ数 $L=3$ であり、各カテゴリ内の像の方向数を $M$ とすると、全辞書画像数は $3M$ である。例えば、3種類のカテゴリのそれぞれに属することがわかっている顔画像を含む多数のテスト画像について図1で説明した照合を行い、各テスト画像と各カテゴリの全ての辞書画像のそれぞれとの最大類似度 $r^i(n, X_{\max}, Y_{\max})$ をそれぞれ求め、全てのテスト画像について得られた類似度を、照合カテゴリ $i=1, 2, 3$ のそれぞれ毎に分類し、各カテゴリ $i$ 毎に類似度に対するテスト画像の数の分布(頻度)を求めると、例えば図10A, 10B, 10Cに示すような結果が得られる。何れのカテゴリの場合も、それぞれ類似度 $r^1_{th}, r^2_{th}, r^3_{th}$ より大の範囲と小の範囲にピークを有する2つの山が生じる。このことは、例えば入力テスト画像をカテゴリ $i=1$ (即ち男性的な顔)と照合した場合、類似度が閾値 $r^1_{th}$ より大であったならば入力テスト画像の顔は男性的であると判定でき、閾値 $r^1_{th}$ より小さければそれ以外、即ち女性的又は中性的顔であると判定できる。従って一般にカテゴリ $i$ との照合において類似度 $r^i(n, X_{\max}, Y_{\max}) > r^i_{th}$ を満足するとき、入力テスト画像はカテゴリ $i$ に属すると判定でき、それぞれのカテ



グリー*i*についての閾値 $r_{th}^i$ と比較することにより入力テスト画像がどのカテゴリーに属するか、あるいは何れにも属しないかが判定できる。

【0025】この実施例によれば、物体の向きによらずに、その物体が予め登録しておいたどのカテゴリーに属するかを識別することができる。図9の実施例の特殊な場合として、入力テスト画像中に人物像があるか否かだけを判定する場合には、辞書画像のカテゴリーとしては1つとし、例えば男性、女性等についての多数の顔画像をそれぞれM方向で取り込み、同じ方向の顔画像を平均化した平均化顔画像をM方向のそれぞれについて求め辞書画像として用意しておく。この様な辞書画像を使って人物無し及び有りの多数のテスト画像を辞書画像と照合すると、図10Dに示すように人物を含むテスト画像の類似度の分布と人物を含まないテスト画像の類似度分布は互いに分離された山となるので、図10Dのように類似度の閾値 $r$ をそれらの山の間に設定することにより、任意の入力テスト画像に人物が含まれているか否かを判定することができる。この場合、図1における辞書画像生成部100に設けられている特徴点検出部103を省略してもよい。あるいは、図9の実施例において辞書画像を生成する場合に、図5A、5Bで説明したように、各カテゴリー内のM方向の辞書画像の隣接する方向角が10度以上、40度未満になるように物体に対する撮像方向をそれぞれ選択することにより、精度の高い照合判定が可能である。

【0026】次に、図6の実施例で示した部分照合の技術と図9の実施例で示したカテゴリーの判定技術を組み合わせ、例えばテスト画像中の人物像が眼鏡を掛けているか否かを判定する場合のように、部分画像を部分辞書画像と照合してその部分画像のカテゴリーを判定するように構成した実施例を図11に示す。図11において、辞書画像生成部100、テスト画像入力部105、照合処理部106、及び物体領域検出部109は図1の対応するものとそれぞれ同じ動作をし、特徴点位置決定部110、部分辞書画像記憶部501、部分照合処理部502、及び特徴点位置高精度決定部505は図6の対応するものとそれぞれ同じ動作をする。

【0027】特徴点位置高精度決定部505は、図6の実施例と同様にテスト画像において辞書画像と最も類似するときの照合位置 $(X_{max}, Y_{max})$ と、そのときの辞書画像番号 $n$ を判定処理部902に出力する。特徴点位置高精度決定部505はまた、図6の実施例と同様にして特徴点 $f$ について、部分辞書画像とテスト画像との類似度が最大となるときの照合位置 $(X_{fmax}, Y_{fmax})$ と部分辞書画像 $I_{Pn}$ とを判定処理部902に出力する。

【0028】判定処理部902は、物体領域検出部109より受けた照合位置 $(X_{max}, Y_{max})$ での $n$ 番目の辞書画像との類似度 $r(n, X_{max}, Y_{max})$ と、特徴点位置高精度決定部505より受けた、位置 $(X_{fmax}, Y_{fmax})$ での部分辞書画像 $I$

$Pn$ との類似度 $r_f(n, X_{fmax}, Y_{fmax})$ とから、テスト画像における物体が辞書画像の物体のどのカテゴリーに属するかを判別し、判別結果を出力する。

【0029】判定処理部902の処理を、テスト画像中の物体が人物のときに、眼鏡をかけていない人物(カテゴリー  $i=1$ )、眼鏡をかけた人物(カテゴリー  $i=2$ )、サングラスをかけた人物(カテゴリー  $i=3$ )のいずれであるかを判別する場合を例に説明する。この場合、辞書画像生成部100内の辞書画像記憶部104(図1参照)に記憶される辞書画像はM方向、3カテゴリーの顔画像である。

【0030】3種類のカテゴリーのそれぞれに属することが予め分かっている多数のテスト画像中の顔画像と眼鏡なしのカテゴリー( $i=1$ )の辞書画像との類似度 $r^1(n, X_{max}, Y_{max})$ の頻度分布を図12Aに示す。この場合、眼鏡無し、眼鏡有り、サングラス有りのカテゴリーがそれぞれ形成する頻度分布の3つの山のピークにおける類似度は、この順に小さくなり、これらの隣接する山の裾は互いに交差している。そのため図12Aに示すように、類似度の閾値 $r_{th}^1$ をカテゴリー $i=1$ の山とカテゴリー $i=3$ の山の間に選択しても、類似度 $r^1(n, X_{max}, Y_{max})$ が閾値 $r_{th}^1$ より大の場合、眼鏡無しの人物または眼鏡有りの人物のいずれかであると判定できるが、それらのどちらであるとも判定できない。そこで、この実施例においては、部分辞書画像記憶部501にカテゴリー $i=1$ 内の各角方向Mの画像に対応して目の部分領域画像 $I_{Pn}$ が記憶されている。

【0031】この場合に対する既知の眼鏡無しのカテゴリー( $i=1$ )に属する人物画像の目の部分領域に関する類似度 $r_f^1(n, X_{fmax}, Y_{fmax})$ の頻度分布を図12Bに示す。同図より、 $r_f^1(n, X_{fmax}, Y_{fmax}) > r_{fth}^1$ のときは、眼鏡無しの人物であると判別すればよいことがわかる。また、 $r^1(n, X_{max}, Y_{max}) \leq r_{fth}^1$ のときは、眼鏡有りの人物またはサングラス有りの人物のいずれかであると識別できる。この場合に対する眼鏡有りの人物のカテゴリーの目の部分領域に関する類似度 $r_f^2(n, X_{fmax}, Y_{fmax})$ の頻度分布を図12Cに示す。同図より、類似度が $r_f^2(n, X_{fmax}, Y_{fmax}) > r_{fth}^2$ のときは、眼鏡有りの人物であると判別すればよいことがわかる。

【0032】図11の実施例において、テスト画像中の物体が眼鏡をかけていない人物であるかどうかを識別する場合には、辞書画像としてカテゴリー数 $L=1$ で、眼鏡をかけていないM方向の顔画像を辞書画像として辞書画像生成部100に用意する。この辞書画像を使ってテスト画像と照合すると、

- (1) テスト画像に顔が含まれていないケース、
- (2) テスト画像に眼鏡をかけた人物を含むケース
- (3) テスト画像に眼鏡をかけていない人物を含むケース

のそれぞれの類似度 $r(n, X_{max}, Y_{max})$ の頻度分布は例えば

図12Dに示すようになる。判定処理部902は物体領域検出部109から出力された類似度 $r(n, X_{\max}, Y_{\max})$ に基づいて図12Dから、 $r(n, X_{\max}, Y_{\max}) > r_{th1}$ のときは、上記ケース(2)または(3)のいずれかであることと識別できる。

【0033】次に、最大類似度を与える照合位置 $(X_{\max}, Y_{\max})$ において前述と同様に特徴点位置を特徴点位置決定部110で決定し、部分照合処理部502で特徴点を含む部分領域と部分辞書画像を照合すると、ケース(2)と(3)のそれぞれに対する部分領域類似度 $r_f(n, X_{f\max}, Y_{f\max})$ の頻度分布は例えば図12Eようになる。そこで判定処理部902は部分領域類似度 $r_f(n, X_{f\max}, Y_{f\max})$ を使って図12Eより、 $r_f(n, X_{f\max}, Y_{f\max}) > r_{th2}$ のときは、眼鏡なしの人物であると判別することができる。

【0034】以上説明した通り、図11の実施例によれば、物体領域全体で照合した場合には差異の小さい別のカテゴリの物体でも、部分領域の照合を行うことにより正しく識別することができる。さらに、図11の実施例において、対象とするカテゴリを論理的に統合することにより、より一般的な識別装置に拡張することができる。例えば、対象とするカテゴリを前述のように $i=1$ :男性顔、 $i=2$ :女性顔、 $i=3$ :サングラスをかけた人物とした場合に、カテゴリ $i=1$ に属するか否かを識別する装置と $i=2$ に属するか否かを識別する装置を組み合わせれば、人物顔( $i=1$  OR  $i=2$ )を識別することができる。また、 $i=1$ に属するか否かを識別する装置と $i=3$ に属するか否かを識別する装置を組み合わせれば、男性でサングラスをかけた人物( $i=1$  AND  $i=3$ )を識別することができる。

【0035】この様に、この実施例によれば、物体領域全体で照合した場合には差異の小さい別のカテゴリの物体でも、部分領域の照合を行うことにより正しく識別することができる。図13は図1の実施例の変形実施例であり、図1に対応するブロックには同じ番号を付けて示して有る。辞書生成部100において図1の場合と同様に、例えば辞書画像入力部101から取り込んだ $L$ 個の各カテゴリの何れかに属する $P$ 人の $M$ 方向の顔画像を顔領域切り出し部102で所望の大きさに切り出す。切り出された顔画像の目と口の特徴点の位置 $(x_f, y_f)$ を特徴点検出部103で検出し、それぞれのカテゴリについて各方向毎に目と口の位置 $(x_f, y_f)$ を用いて顔領域の大きさと位置を合わせてから平均する。この様にして $L \times M$ 個の平均画像を作成する。

【0036】この実施例では更に、モザイク処理部121が設けられ、これによって各平均画像を $v \times w$ のブロックに分割し、それぞれのブロック内の画素濃度を平均化する。この処理をモザイク処理と呼び、モザイク処理された画像をモザイク画像と呼ぶ。1つのモザイク画像の全ブロック( $v \times w$ 個)の濃度値の列を特徴列と呼び、

$v \times w$ 次元のベクトルを構成している。図14は、平均画像を $12 \times 12$ のブロックに分割した例を示している。この場合、特徴列は $144$ 次元のベクトルである。各モザイク画像の特徴列は辞書生成部122に与えられる。

【0037】辞書生成部122は、与えられた $L$ 個の各カテゴリについて $M$ 個の特徴列 $f_{i,m}(m=1,2,\dots,M)$ を受けると、カテゴリを示す個別コード $i$ と $M$ 個の特徴列を対応付けた、図15に示すようなテーブルを作成し、辞書記憶部123に書き込む。テスト画像入力部105は取り込んだ画像をテスト画像として照合処理部106に出力する。

【0038】照合処理部106は図1と同様の照合領域切り出し部107Aと類似度算出部107と位置シフト部108に加えて、モザイク処理部124が設けられている。モザイク処理部124の処理は辞書生成部100内のモザイク処理部121と同様のモザイク処理を行う。照合領域切り出し部107Aはテスト画像入力部105からテスト画像を受けると、位置シフト部108から与えられた照合位置の初期値 $(X_0, Y_0)$ を基準に所定の大きさの照合領域の画像を切り出しモザイク処理部124に与える。モザイク処理部124は与えられた照合領域画像をモザイク処理し、処理された画像の特徴列を類似度算出部107に与える。類似度算出部107は辞書記憶部123に辞書の読み出し要求を出し、 $N=L \times M$ 個の辞書特徴列(即ち辞書画像)を順次受け、画像特徴列と各辞書特徴列との類似度 $r(n, X_0, Y_0)$ をすべて算出し、それらの類似度を物体領域検出部109に出力するとともに、照合位置更新命令を位置シフト部108に与える。位置シフト部108は図1の場合と同様に照合位置を更新して新しい照合位置 $(X, Y)$ を出力する。照合領域切り出し部107Aは更新された照合位置でテスト画像から照合領域を切り出し、モザイク処理部124に与える。モザイク処理部124はその照合領域画像のモザイク処理を行い特徴列を算出し類似度算出部107に与え、テスト画像におけるその照合領域と辞書画像との類似度 $r(n, X, Y)$ を算出して物体領域検出部109に出力するとともに、照合位置更新命令を位置シフト部108に与える。上述の処理を繰り返し、照合位置 $(X, Y)$ が予め設定した値であれば、照合処理部106の処理を終了し、物体領域検出部109の処理に進む。類似度 $r$ は、例えば、2つの特徴列の間のユークリッド距離の逆数により算出すればよい。

【0039】物体領域検出部109は図1の場合と同様に、受け取った類似度 $r(n, X, Y)$ が最大となるときに照合位置 $(X_{\max}, Y_{\max})$ と辞書画像番号 $n$ を検出する。この様に、図13の実施例によれば、照合領域を $v \times w = q$ 個のブロックにしてモザイク処理することにより、対象領域の位置ズレや大きさの違いやボケに対し強くなり、又全体としての計算量が減少するため処理を高速にでき、更に辞書画像を記憶するための記憶容量を削減することができる。

【0040】次に、図13の実施例を更に変形した実施例を図16に示す。辞書生成部100の処理は図13の場合と同じであり、説明を省略する。この実施例においては更に、背景像保持部125、物体領域抽出部126、物体計測部127、ブロックサイズ限定部128、及びブロックサイズ変更部129が設けられる。背景画像保持部125は、テスト画像Iを得る例えば一定時間前に同じ方向でテスト画像入力部105によって撮影した画像を背景像として保持する。物体領域抽出部126はその処理を図17に概念的に示すように、テスト画像入力部105からのテスト画像 $I_T$ と背景画像保持部125からの背景画像 $I_B$ の差分処理を行って差分画像 $I_{df}$ を得て、その差分画像の全ての画素値を2値化処理して2値化画像 $I_{db}$ を生成する。図17では人物像領域の画素が1、その背景の斜線領域の画素が0であるとする。この2値画像 $I$ の1についてX軸及びY軸への射影をそれぞれ求めるとグラフ(a),(b)で示すようなそれぞれの座標上の画素数の分布が得られる。この分布(a),(b)の形状から人物像の頭部の範囲、従ってその大きさ $D_x$ ,  $D_y$ が図に示すように求められる。

【0041】ブロックサイズ限定部128は、前記 $D_x$ ,  $D_y$ をもとに、実際の認識対象物体の大きさが横 $D_x - dx$ から $D_x + dx$ 、縦 $D_y - dy$ から $D_y + dy$ の範囲にあると仮定して、 $v \times w$ ブロックのサイズの可変範囲を次のように決める。

横： $(D_x - dx)/v$ 画素から $(D_x + dx)/v$ 画素の範囲  
縦： $(D_y - dy)/w$ 画素から $(D_y + dy)/w$ 画素の範囲  
但し、割れきれない場合は四捨五入等で整数値にする。

【0042】類似度算出部107は、テスト画像入力部105からテスト画像を受けると、辞書記憶部123に辞書画像(辞書特徴列)の読み出し要求を出し、辞書特徴列を受け取る。ブロックサイズ変更部129は、ブロックサイズ限定部128で限定した範囲でブロックサイズを初期値にfixし、モザイク処理部124、類似度算出部107、位置シフト部108は図13で説明したと同様の処理を行い、全ての照合位置においてモザイク処理と類似度算出処理を終えたらブロックサイズ変更部129でブロックサイズを一定幅ずつ変更して同様の処理を繰り返す。

【0043】認識対象物体の大体の大きさを検出することにより、ブロックサイズの可変範囲を限定することができ、処理時間を短縮することができる。図18は各辞書画像の特徴点が予め決めた位置関係となるように辞書用画像を幾何学的正規化処理し、テスト画像についても特徴点の位置関係が対応する正規化辞書画像の特徴点位置と一致するように幾何学的正規化処理をする事により、照合精度を高めるようにした実施例を示し、特徴点として、左右の目領域の重心を $P_1$ と $P_2$ 、口領域の重心を $P_3$ として抽出する場合を示す。この実施例では、位置関係情報記憶

部131、幾何学的正規化処理部132及び認識用辞書画像記憶部133から成る認識用辞書生成部130が設けられる。一方、テスト画像入力側にも更に幾何学的正規化処理部134、類似度算出部135及び判定処理部136が設けられる。

【0044】認識用辞書生成部130の位置関係記憶部131には、予め決めた基準となる方向から参照物体を撮影した場合の画像における特徴点の位置関係を記憶している。この実施例では図19Aに示す例えば正面顔における3つの特徴点 $P_1, P_2, P_3$ の次の位置関係を記憶しておく。

- (1)  $P_1$ と $P_2$ を結ぶ直線は水平方向
- (2)  $P_1$ と $P_2$ の中点 $P_0$ と $P_3$ を結ぶ直線は垂直方向
- (3)  $P_1$ と $P_2$ の距離
- (4)  $P_0$ と $P_3$ の距離

幾何学的正規化処理部132は、位置関係情報記憶部131に読み出し要求を出し、特徴点 $P_1, P_2, P_3$ の位置関係を受け取る。次に幾何学的正規化処理部132は、辞書画像生成部100内の辞書画像記憶部104から読み出された辞書画像 $I_{Dn}$ とそれに対応する特徴点の位置( $x_f, y_f$ ),  $f=1, 2, 3$ を受け、それらの特徴点の位置関係が位置関係情報記憶部131から受けた特徴点 $P_1, P_2, P_3$ の位置関係と一致するような幾何学的変換、例えば、アフィン変換を受け取った辞書画像 $I_{Dn}$ に施し、変換後の画像データから切り出した照合領域を認識用辞書画像 $I'_{Dn}$ として認識用辞書画像記憶部133に記憶する。この処理を辞書画像生成部100から受けた全ての辞書画像 $I_{Dn}$ ,  $n=1, \dots, N$ について実行する。照合領域の切り出しは、 $P_1$ と $P_2$ 間の距離の定数倍を幅、 $P_0$ と $P_3$ 間の距離の定数倍を長さとする矩形領域を切り出せばよい。図19Bに、求めた認識用辞書画像の一例を示す。

【0045】テスト画像がテスト画像入力部105から入力されると、図1の実施例と同様に辞書画像生成部100からの辞書画像と照合処理部106で照合して、その結果に基づき物体領域検出部109でテスト画像中の物体領域(ここでは顔領域)を検出し、更に特徴点位置決定部110で目及び口の特徴点位置( $X_f, Y_f$ ),  $f=1, 2, 3$ を決定する。幾何学的正規化処理部134は検出された特徴点( $X_f, Y_f$ )の位置関係が位置関係情報記憶部131に保持されている特徴点 $P_1, P_2, P_3$ の位置関係と一致するように顔領域画像をアフィン変換し、類似度算出部135に出力する。類似度算出部135は認識用辞書画像記憶部133から認識用辞書画像 $I'_{Dn}$ を順次読みだし、テスト画像中の変換された顔領域画像との類似度を算出し、判別処理部136に出力する。類似度としては、例えば2つの照合領域の画像の相関係数を用いればよい。判別処理部136では類似度が最大の辞書画像が代表するカテゴリーを判別結果として出力する。

【0046】この様に図18の実施例によれば、対象物体領域から特徴点を抽出し、抽出した特徴点の位置関係

が、予め記憶しておいた基準となる方向から参照物体を撮影したときの画像から抽出した特徴点の位置関係と一致するように、取り込んだテスト画像における物体領域を幾何学的に変換し、変換後の画像の類似度に基づいて物体領域を抽出する。この様に画像を幾何学的変換するので、撮影する方向により画像における物体の形状の大きさ、傾き等が変わる場合でも、辞書画像作成する際に同一参照物体を異なる方向から撮影する回数を低減させることができるし、テスト画像における物体の形状や大きさが変わる場合でも、物体検出処理に要する計算量を削減させ、かつ検出精度を向上させることができる。

【0047】図20は図1の実施例において、テスト画像の照合領域と辞書画像の明るさをそれぞれ正規化することにより照合精度を高めるように構成した実施例を示す。なお、特徴点位置決定部110は省略される。この実施例では、辞書画像生成部100と照合処理部106にそれぞれ明るさ正規化部137, 138を設け、特徴点検出部103及び照合領域切り出し部107Aからそれぞれ受けた画像の照合領域の明るさの正規化を行う。従って辞書画像記憶部104には明るさが正規化された辞書画像が書き込まれ、また類似度算出部107も明るさが正規化された照合領域画像が与えられる。明るさの正規化は、例えば、照合領域の濃度値の平均と分散が予め設定しておいた値となるように、各画素の濃度値を変換することにより実現できる。別の正規化方法としては、照合領域の濃度値の最小値と最大値をそれぞれ予め設定した値に変換し途中の値も線形或は非線形変換することによっても実現できる。他の部分は図1の実施例と同様である。

【0048】図20の実施例と同様に、図9の実施例における辞書画像生成部100内と照合処理部106内にそれぞれ明るさ正規化部137, 138を設けてもよい。この様に、図20の実施例によれば、物体領域内の各画素の濃度の平均と分散が予め決めた値になるように各画素濃度を変換し、変換後の画像から計算した類似度により物体領域を検出する。このような明るさの正規化を行うことで、撮影する方向や他の条件により、テスト画像における物体の明るさが大きく変わっても、辞書画像との照合精度の低下を防ぐことができる。

【0049】以上説明した様に、本発明は、対象となるカテゴリに属する物体をあらかじめ所定のM方向からカメラで撮影し、得られた画像から該方向と該カテゴリを代表するN個の辞書画像を生成し、記憶しておき、テスト画像が与えられると、該テスト画像の各位置において、N個の辞書画像との照合をとり、最も類似するときのテスト画像の位置を検出することにより画像中の物体領域を抽出し、また、そのときの類似度から該物体が対象となるカテゴリに属するか否かを識別するので、物体の撮影方向によらずに、画像中より該物体領域を高速な処理で安定に抽出でき、さらに、抽出された物体があらかじめ定めたカテゴリ（人物顔、子供顔、成人女

性、成人男性）に属しているか否かを識別できる。

【0050】本発明の応用として、銀行などでの出入口の自動ドアの圧力センサ出力や現金自動支払機の操作開始をトリガーにして取り込んだ人物画像を、本発明の物体識別装置のテスト画像とした場合には、対象となるカテゴリを人物顔として動作すれば、通常の顔（客）とサングラスやマスクで顔を隠した人物（不審者）を識別できるので、不審者を検出するシステムが実現できる。

【0051】また、商店街を通過する人物やショーウィンドーを見ている人物の画像を本発明の識別装置に入力し、対象となるカテゴリとして、男性顔、女性顔、眼鏡をかけた人などを用意して、テスト画像の人物を各カテゴリに分類することにより、どのような人物がどのような物に関心を持っているかを分析する市場調査システムに應用できる。

【図面の簡単な説明】

【図1】この発明の第1の実施例を示す機能ブロック図。

【図2】図1の辞書画像記憶部104に記憶されている画像の例を示す図。

【図3】図1の照合処理を説明する図。

【図4】撮影方向角度間隔と物体領域抽出成功率の関係を示すグラフ。

【図5】A及びBは撮影方向角を説明するための図。

【図6】この発明の第2の実施例を示す機能ブロック図。

【図7】部分辞書画像の生成を説明する図。

【図8】部分照合処理を説明する図。

【図9】この発明の第3の実施例を示す機能ブロック図。

【図10】A～Dは各カテゴリの画像の類似度分布を示すグラフ。

【図11】この発明の第4の実施例を示す機能ブロック図。

【図12】A～Eは各カテゴリの画像の類似度分布を示すグラフ。

【図13】この発明の第5の実施例を説明する機能ブロック図。

【図14】モザイク処理を説明するための図。

【図15】辞書画像の特徴列を示す表。

【図16】この発明の第6の実施例を示す機能ブロック図。

【図17】画像中の物体の範囲を抽出する処理を示す図。

【図18】この発明の第7の実施例を示す機能ブロック図。

【図19】A, Bは特徴点の関係を示す図。

【図20】この発明の第8の実施例を示す機能ブロック図。

【図1】

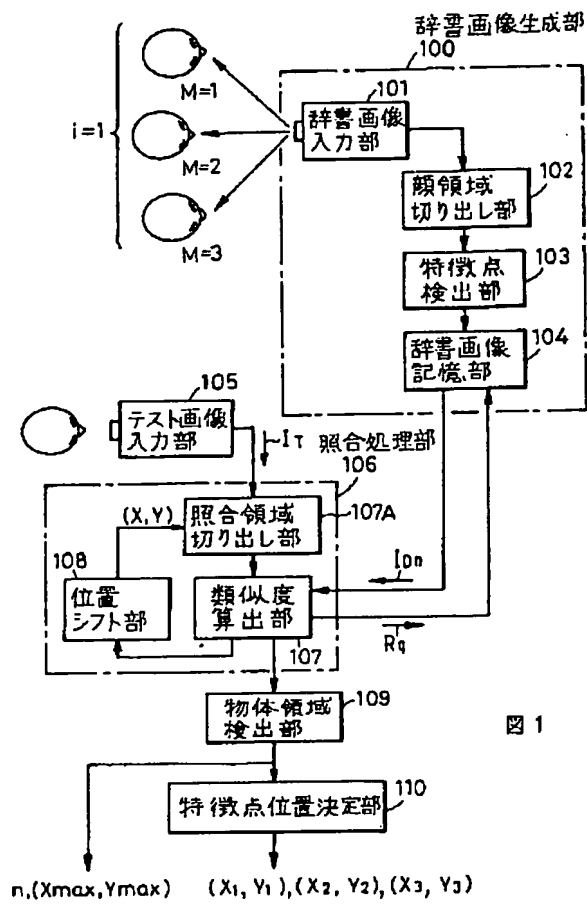


図 1

【図2】

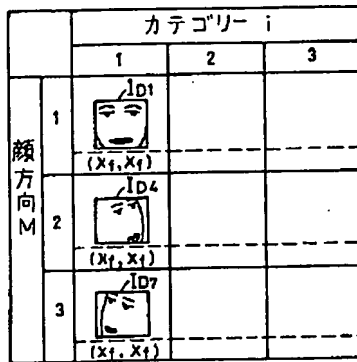


図 2

【図5】

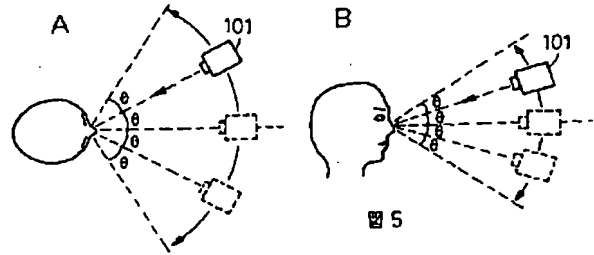


図 5

【図3】

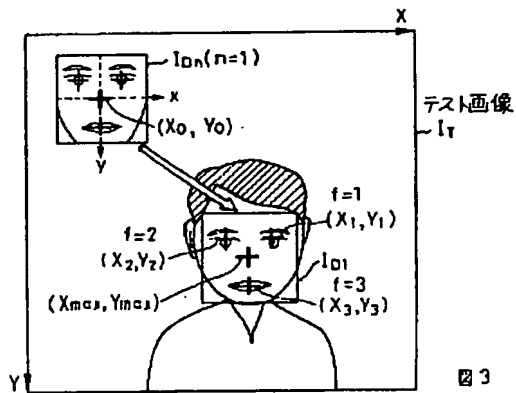


図 3

【図4】

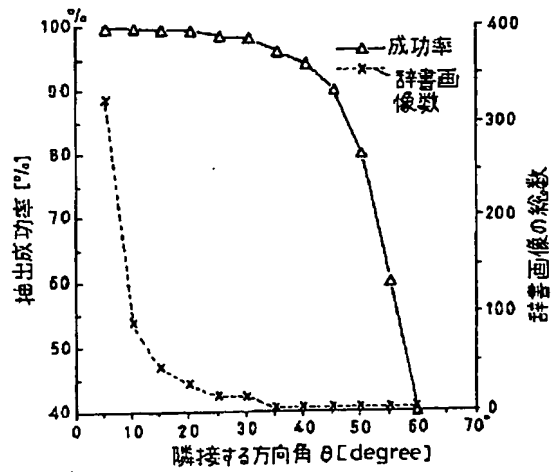
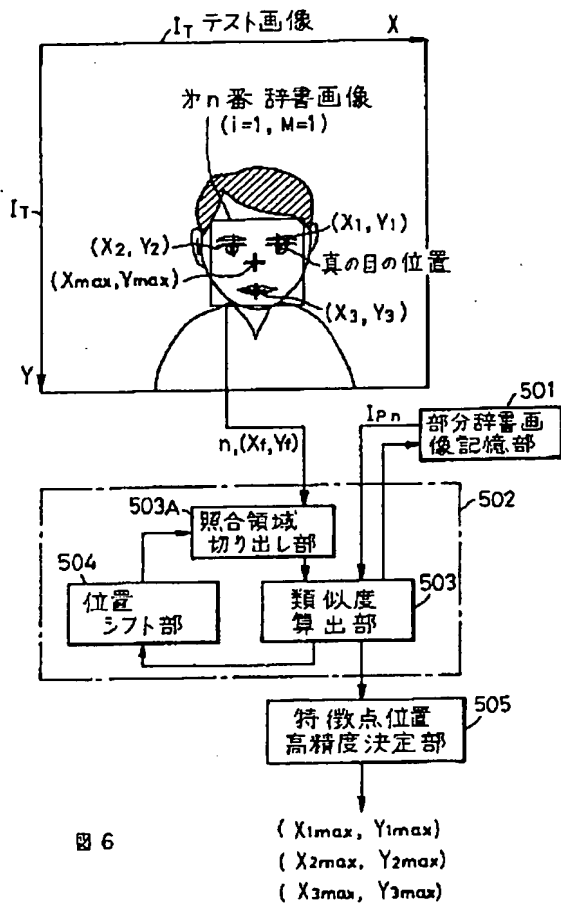
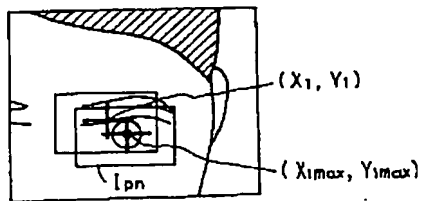


図 4

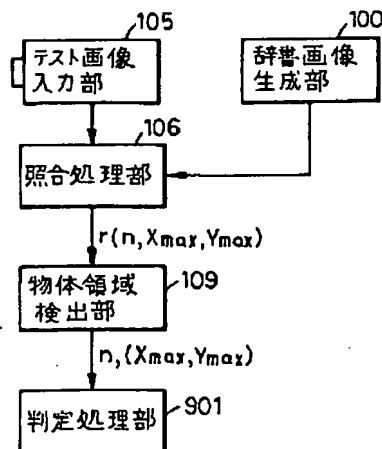
【図6】



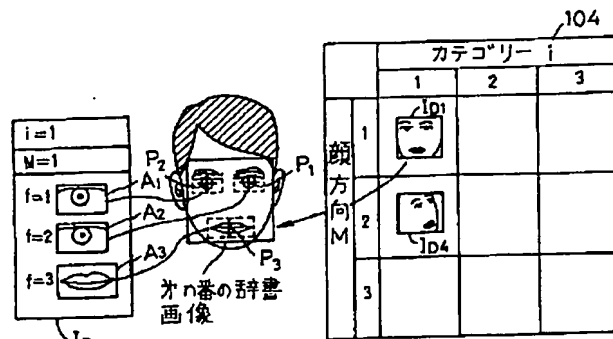
【図8】



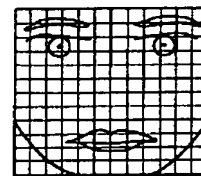
【図9】



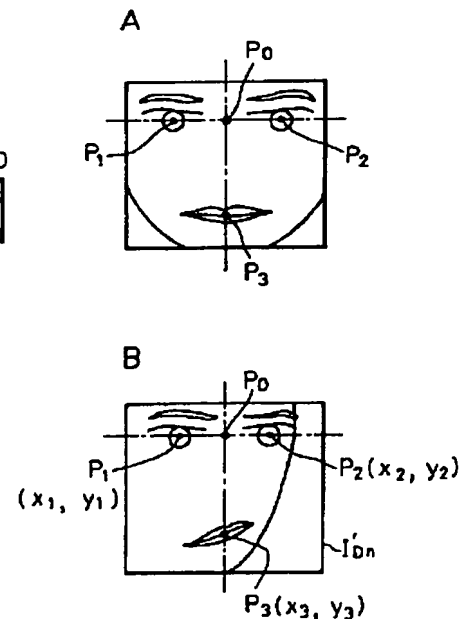
【図7】



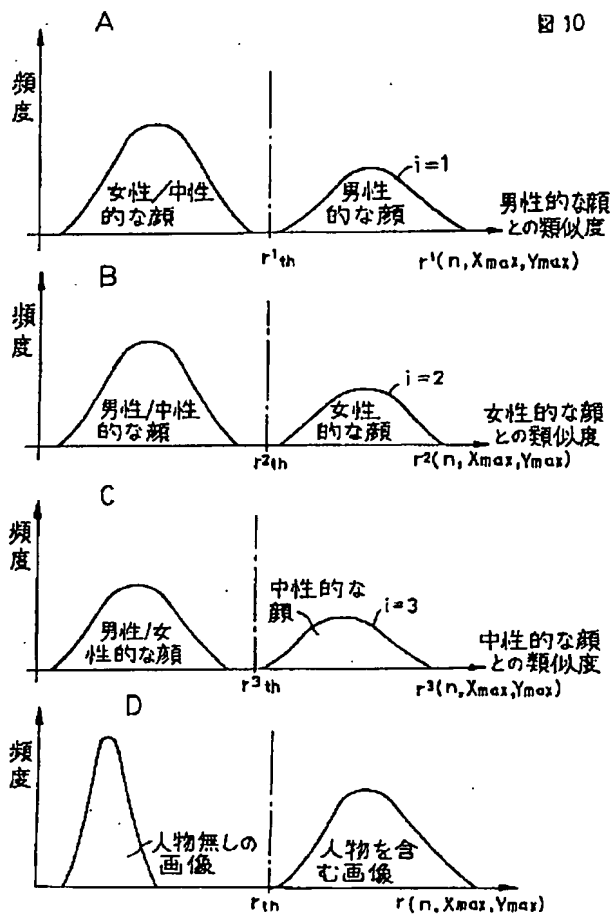
【図14】



【図19】



【図10】



【図15】

カテゴリー	特徴列
1	$f_{1,1}, f_{1,2}, \dots, f_{1,M}$
2	$f_{2,1}, f_{2,2}, \dots, f_{2,M}$
$\vdots$	$\vdots$
$i$	$f_{i,1}, f_{i,2}, \dots, f_{i,M}$
$\vdots$	$\vdots$
L	$f_{L,1}, f_{L,2}, \dots, f_{L,M}$

図15

【図11】

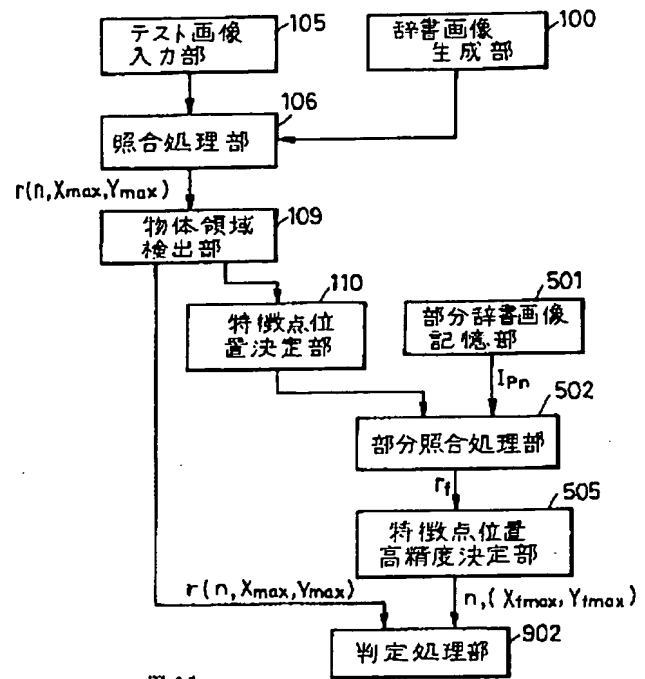


図11

【図13】

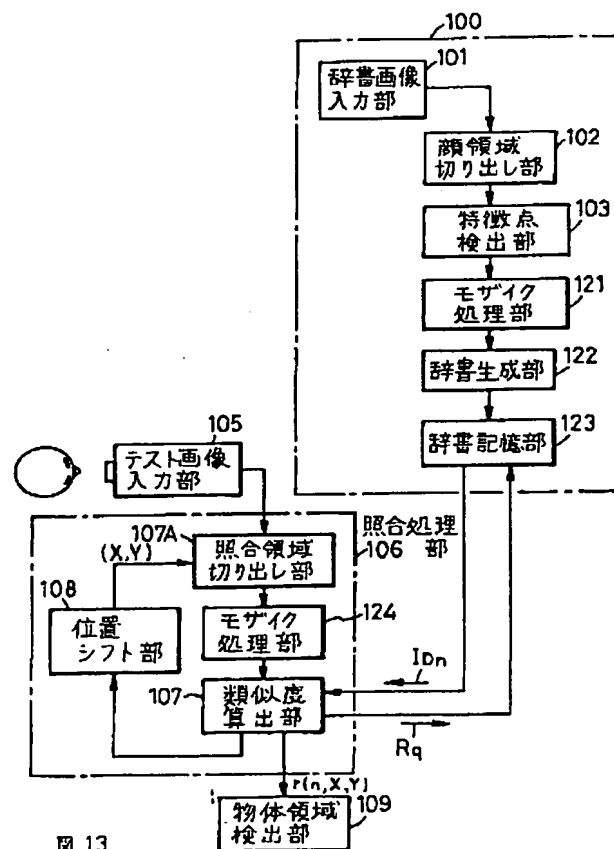


図13

【図12】

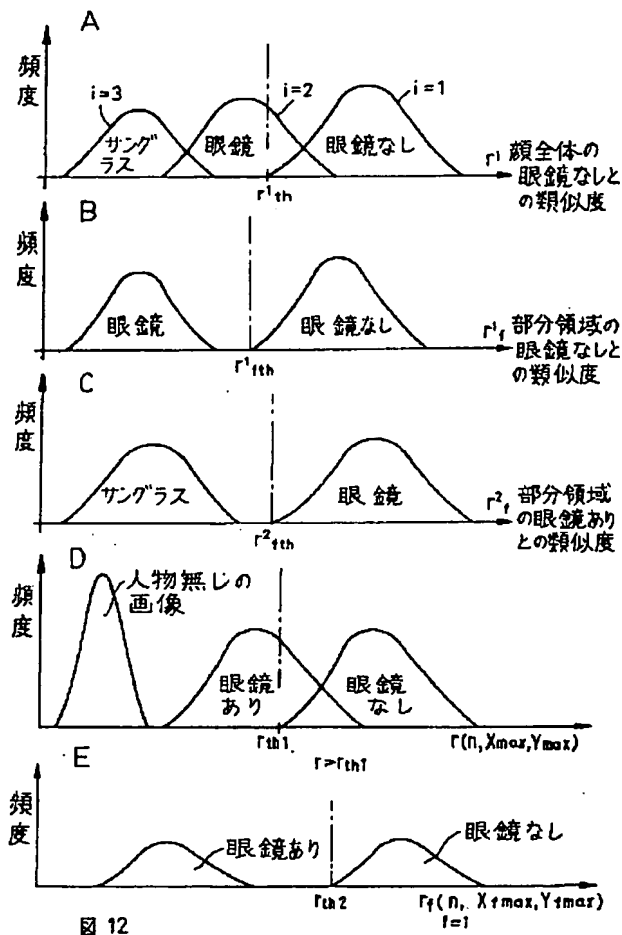


図 12

【図16】

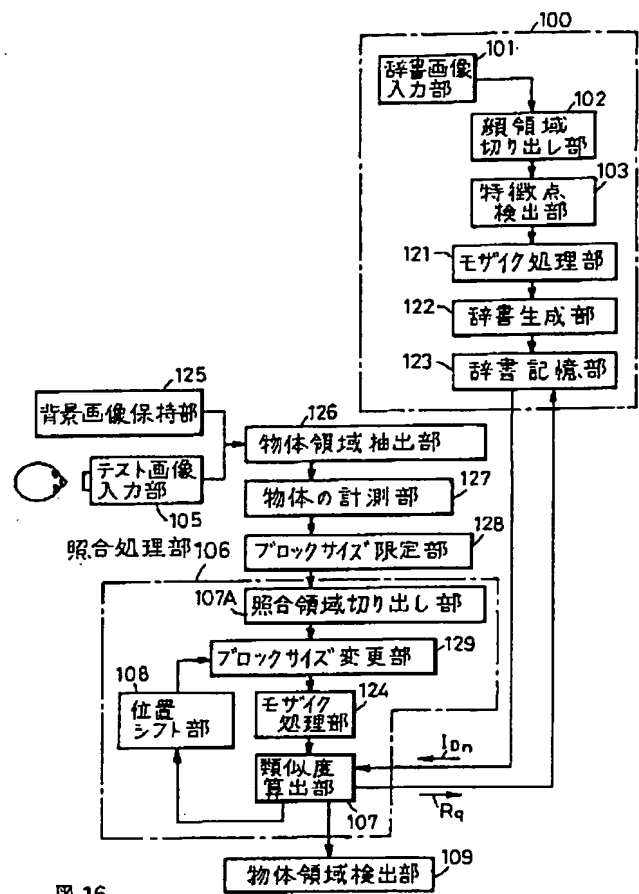


図 16



【図17】

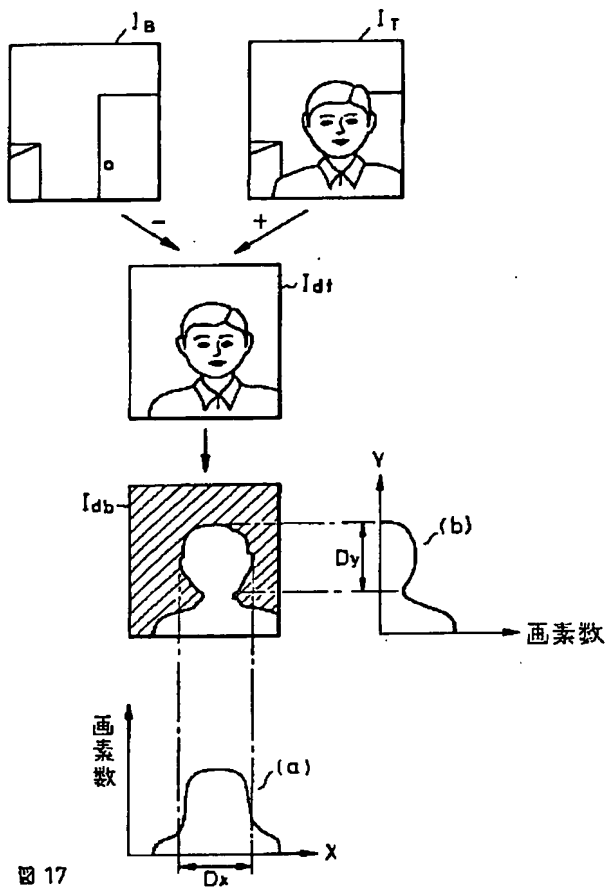


図 17

【図18】

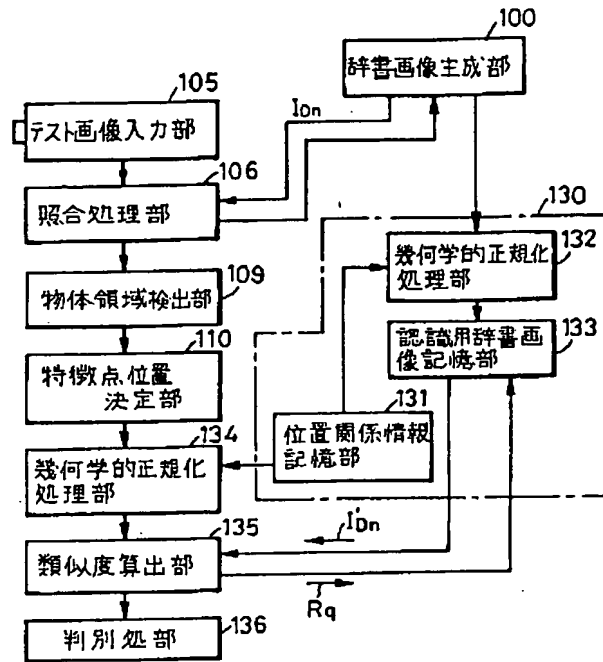


図 18

【図20】

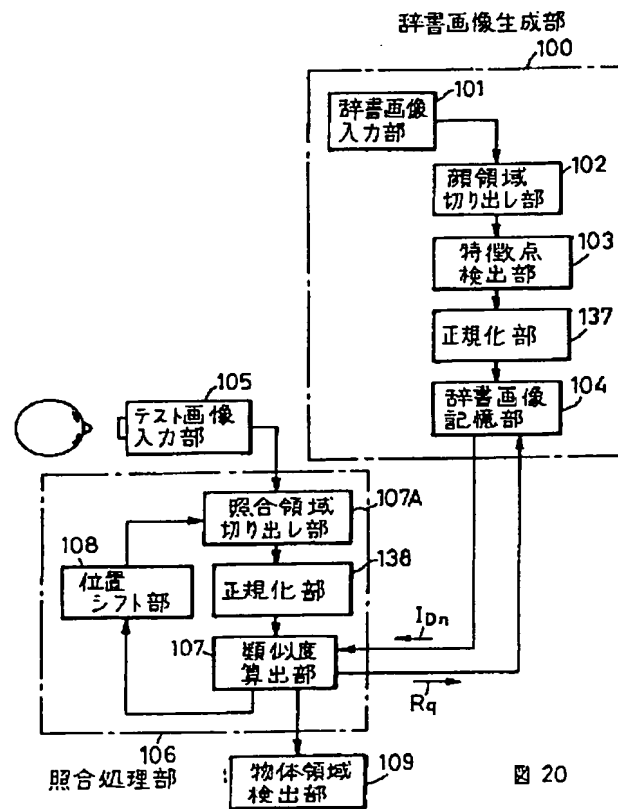


図 20

フロントページの続き

(72)発明者 石井 健一郎  
東京都千代田区内幸町1丁目1番6号 日  
本電信電話株式会社内

(72)発明者 磯 俊樹  
東京都千代田区内幸町1丁目1番6号 日  
本電信電話株式会社内

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**